# NON-PARAMETRIC REGRESSION FOR PATCH-BASED FLUORESCENCE MICROSCOPY IMAGE SEQUENCE DENOISING

*J. Boulanger*[1], *J.-B. Sibarita*[1] *, C. Kervrann*[2,3]*, P. Bouthemy*[3]

(1) UMR 144 Insitut Curie/CNRS 26 rue d'Ulm, F-75005 Paris
(2) INRA, Unité de Biométrie Domaine de Vilvert F-78352 Jouy-en-Josas
(3) IRISA/INRA, Campus de Beaulieu, F-35042 Rennes

## ABSTRACT

We present a non-parametric regression method for denoising fluorescence video-microscopy volume sequences. The designed method aims at using the 3D+t information in order to restore acquired data contaminated by Poisson and Gaussian noise. We propose to use a variance stabilization transform to deal with the combination of Poisson and Gaussian noise. Consequently, we further propose an adaptive patch-based framework able to preserve space-time discontinuities and reduce significantly noise level using the 3D+t space-time context. This approach lead to an algorithm whose parameters are calibrated and then ready for intensive use. The performance of the proposed method are then demonstrated on both synthetic and real image sequences using quantitative as well as qualitative criteria.

*Index Terms*— Video-microscopy, image sequence denoising, patch-based approach, Poisson noise, adaptive estimation.

## 1. INTRODUCTION

Video-microscopy is an effective tool for the analysis of dynamical phenomena occurring inside living cells. Combined with fluorescence, both confocal microscopy and wield-field microscopy allow the localization of proteins of interest in 3D space and time. Unfortunately, three factors limit the number of photons that can be used. First, photo-bleaching prevent the use of high illumination intensities. Second, photo-toxicity perturbes the cell cycle up to cell death. Third, observed fast kinetics require high acquisition rates.

By improving the signal to noise ratio, denoising allows to reduce the amount of light needed to record images and thus higher frame rate will make possible to observe faster phenomena without perturbing the activity of living cells. Moreover, besides of better visualization of the acquired data, some analysis like spot detection and tracking can be strongly helped by a denoising procedure.

Few methods use the temporal coherence for restoration of video-microscopy image series. However, for video application, a review of image sequence denoising methods can be found in [1]. Kuznetsov et *al.* proposed to use a temporal Kalman-Bucy filter to improve the quality of video-microscopy image sequences [2]. Estimating the motion of small objects moving with high velocity is a difficult task, alternatively methods extended from still image denoising domain do not rely explicitly on motion estimation. Thus, wavelet shrinkage [3, 4], Wiener filtering [5] or PDE-based methods [6] has been applied to image sequence denoising.

Recently an extension of the non-local mean filter has been proposed for image sequences [7]. An improvement proposed in [8] consider a collection of similar patch as 3D arrays. An unitary transform is then applied and a hard-threshold is used to attenuate noise.

In an other framework, also related to the non-local mean filter, considers that images admit a sparse approximation over a dictionary of atoms which can be a DCT basis or a set of image patches [9]. This approximation can be formalized as the minimization of a functional involving a data term and a penalty related to the size of the dictionary.

Low intensities in video-microscopy lead to Poisson statistics. On the other hand, electronic Gaussian white noise remains, despite the use of cooling system for the sensors, resulting in a combination of Poisson noise and Gaussian noise. In the past, several approaches have been introduced to deal with signal dependent noises. Maximum likelihood estimator can be derived for Poisson statistic (e.g. [10]). However, the extension to the combination on Poisson and Gaussian is not trivial. One can also apply a variance stabilization transform [11]. A generalized Anscombe transform for Poisson and Gaussian noise can be found in [12] and is appropriate when then number of photon is large enough. Finally, on-line adaptation to the local noise statistics is a third way to deal with signal dependent noise (e.g. [13]).

In this paper, we present a space-time patch based adaptive statistical method for video-microscopy image sequence restoration. We propose to take into account the statistical properties of the noise using a realistic modeling. Furthermore, we consider 3D image sequences as sequences of volumes and use the whole 3D+t information to restore noisy data. Our approach involves a patch-based functional leading to the definition of an estimator of the unknown noise-free function and a statistical analysis of the performance of this estimator allows us to adapt locally the neighborhood containing similar patches. Resulting algorithm has only few parameters which are all calibrated.

The remainder of this paper is organized as follow: we first introduce the problem of video-microscopy image sequence denoising and describe the proposed method in Section 2. Finally, we demonstrate in Section 3 the ability of our approach at restoring video-microscopy image sequence using both synthetic and real data.

## 2. PROPOSED METHOD

In this section, we present a general framework for image sequence analysis in fluorescence microscopy. Our study is limited to the restoration of damages or artifacts due to random noise. Accordingly, we do not consider the important issue of correcting the distortions of the signal resulting from the optical system of microscopes (e.g. deconvolution problem).

## 2.1. Problem statement

Acquired images correspond to stacks of 10 to 60 slices with a resolution in the axial direction $z$ of typically 1/3 when compared to the resolution in the lateral direction $x - y$. In our study, the images show GFP-tagged proteins appearing as bright particles of size 5 to 10 pixels moving with speeds ranging from 1 to 10 pixels. The small amount of light collected by sensors and thermal agitation in electronic components leads to a mixed Poisson-Gaussian noise. In what follows, we assume the following linear model

$$Z_i = g_0 N_i + \varepsilon_i, \tag{1}$$

where $Z_i$ is the observation at the space-time location $\mathbf{x}_i \in \mathbb{R}^4$ (also denoted by the index $i$) and $g_0$ represents the gain of the overall electronic system. The number of collected photo-electrons is a random variable assumed to follow a Poisson distribution with parameter $\theta_i$: $N_i \sim \mathcal{P}(\theta_i)$. Finally, the dark current is modeled by a Gaussian white noise of variance $\mathrm{Var}[\varepsilon_i] = \sigma_\varepsilon^2$ and expectation $\mathbb{E}[\varepsilon_i] = m$.

In this paper, we consider the problem of estimating $\theta_i := \theta(\mathbf{x}_i)$ at each point $\mathbf{x}_i$ from noisy data $Z_i$ taken in a space-time neighborhood of $\mathbf{x}_i$.

## 2.2. Stabilization of noise variance

In this paper, we consider that the number of collected photoelectron is large enough and use the generalized Anscombe transform (GAT) [12]:

$$\mathcal{T}_{GA}(Z_i) = \frac{2}{g_0} \sqrt{g_0 Z_i + \frac{3}{8} g_0^2 + \sigma_\varepsilon^2 - g_0 m}. \tag{2}$$

The GAT requires the knowledge of the three parameters $g_0$, $\sigma_\varepsilon^2$ and $m$. These ones can be obtained by calibrating of the acquisition system or can be estimated e.g. by dichotomy. We propose here an approach based on a linear regression in the plane $(\mathbb{E}[Z_i], \mathrm{Var}[Z_i])$. From (1), we have

$$\begin{cases} \mathbb{E}[Z_i] &= g_0 \theta_i + m, \\ \mathrm{Var}[Z_i] &= g_0^2 \theta_i + \sigma_\varepsilon^2. \end{cases} \tag{3}$$

which yields

$$\mathrm{Var}[Z_i] = g_0 \mathbb{E}[Z_i] + \sigma_\varepsilon^2 - g_0 m. \tag{4}$$

A linear regression in the plane $(\mathbb{E}[Z_i], \mathrm{Var}[Z_i])$ provides an estimation of the two parameters: $g_0$ and $e_{DC} = \sigma_\varepsilon^2 - g_0 m$. Accordingly, (2) can be stated as

$$\mathcal{T}_{GA}(Z_i) = \frac{2}{g_0} \sqrt{g_0 Z_i + \frac{3}{8} g_0^2 + e_{DC}}. \tag{5}$$

Now, our idea is to robustly estimate the local mean and the local variance. In order to get independent samples, it is crucial to partition the space-time domain into non-overlapping regions. The size of these regions results from a compromise between the estimators variances and the number of resulting measure points in the plane $(\widehat{\mathbb{E}[Z_i]}, \widehat{\mathrm{Var}[Z_i]})$. In each region, the mean of the signal is estimated using a robust M-estimator while the noise variance is robustly estimated using the "Least Median of Square" estimator defined as $\widehat{\mathrm{Var}[Z_i]} = 1.4826 \, \mathrm{med}_i \, (\mid r_i - \mathrm{med}_j \mid r_j \mid \mid)$. where the pseudo-residuals $r_i$ are computed at each spatial position $(x, y)$ (in the 2D case) as [14]:

$$r_x = \frac{1}{\sqrt{l^2 + l}} \Delta Z(x) \tag{6}$$

where $\Delta Y(x)$ is a space-time Laplacian operator involving $l$ surrounding pixels. Given estimates of the mean and variance, a robust linear regression is applied to derive the target parameters $g_0$ and $e_{DC}$ and apply the GAT to stabilize the variance of the noise.

## 2.3. Patch based space-time estimation

Once that the noise variance has been stabilized, we consider now the following image sequence model:

$$Y_i = u(\mathbf{x}_i) + \eta_i, \tag{7}$$

where $\mathbf{x}_i \in \Omega$ denotes the pixel location in the space-time volume $\Omega \subset \mathbb{R}^4$. The regression function $u_i = u(\mathbf{x}_i)$ is the ideal image to be recovered from observations $Y_i := \mathcal{T}_{GA}(Z_i)$. The errors $\eta_i$ are now assumed to be independent zero-mean Gaussian variables with variance $\sigma_\eta^2$ theoretically equal to 1. Once $u$ has been estimated, the inverse generalized Anscombe transform is applied to recover $\theta_i$ at each spatial position.

Let us define a patch as a sub-region of the space-time domain of size $p \times p \times p$ and consider that each patches $\widehat{\mathbf{u}}_i$ of the space-time domain as a vector that can be expressed as a linear combination of original noisy patch $\mathbf{Y}_j$ lying inside a neighborhood $\Delta(\mathbf{x}_i)$ of the point $x_i \in \Omega$. We define the estimator of the patch $\mathbf{u}_{i,n}$ at iteration $n$ as the following weighted average:

$$\widehat{\mathbf{u}}_{i,n} = \sum_{\mathbf{x}_j \in \Delta(\mathbf{x}_i)} w_{ij,n} \mathbf{Y}_j \tag{8}$$

were $w_{ij,n}$ denotes the weights. We can then estimate the variance of this estimation using:

$$\widehat{\mathbf{v}}_{i,n} = \sum_{\mathbf{x}_j \in \Delta(\mathbf{x}_i)} w_{ij,n}^2 \tag{9}$$

since the variance of the noise has been stabilized to 1. In order to select similar patch we use the following weights:

$$w_{ij,n} = \frac{\exp\left(-\frac{1}{2} \|\widehat{\mathbf{u}}_{i,n-1} - \widehat{\mathbf{u}}_{j,n-1}\|_{\widehat{\boldsymbol{\Sigma}}_{i,n-1}}^2\right)}{\sum_{\mathbf{x}_k \in \Delta(\mathbf{x}_i)} \exp\left(-\frac{1}{2} \|\widehat{\mathbf{u}}_{i,n-1} - \widehat{\mathbf{u}}_{k,n-1}\|_{\widehat{\boldsymbol{\Sigma}}_{i,n-1}}^2\right)}, \tag{10}$$

where the distance between two regularized patch $\widehat{\mathbf{u}}_{i,n-1}$ and $\widehat{\mathbf{u}}_{j,n-1}$ is computed as:

$$\|\widehat{\mathbf{u}}_i - \widehat{\mathbf{u}}_j\|_{\widehat{\boldsymbol{\Sigma}}_i}^2 = (\widehat{\mathbf{u}}_i - \widehat{\mathbf{u}}_j)^T \widehat{\boldsymbol{\Sigma}}_i^{-1} (\widehat{\mathbf{u}}_i - \widehat{\mathbf{u}}_j) \tag{11}$$

Besides, the symmetric scaling matrix $\widehat{\boldsymbol{\Sigma}}_i$ is related to the variance obtained at the previous iteration:

$$\widehat{\boldsymbol{\Sigma}}_{i,n-1} = \frac{\lambda}{2} \mathrm{diag}(\widehat{\mathbf{v}}_{i,n-1} + \widehat{\mathbf{v}}_{j,n-1}). \tag{12}$$

where the parameter $\lambda$ is a $\alpha$-quantile of the $\chi_{p^3}^2$ distribution. Finally, the set of overlapping patches $\{\widehat{\mathbf{u}}_k : \|\mathbf{x}_k - \mathbf{x}_i\|_1 < p\}$ provide several estimates at point $\mathbf{x}_i$ which can be simply fused by taking the average. This also allow to speed up the algorithm by sub-sampling the image domain. Experiments show that using a subsampling factor of $1/(p+1)$ give the same result than a non-vectorized version of the algorithm. By using no sub-sampling, the quality of the denoising is improved.

## 2.4. Neighborhood size selection

The estimator (8) is based on the approximation of the central patch by a set of nearby patches. The performance of the estimator is then related to the radius of this neighborhood and can vary at each point of the image sequence according to image contents.

In order to optimally estimate the neighborhood size, we analyze the performance of the estimator and consider the usual local $L_2$ risk defined as $\mathcal{R}(\widehat{u}_i, u(\mathbf{x}_i)) = \mathbb{E}[(\widehat{u}_i - u(\mathbf{x}_i))^2]$. where $u(\mathbf{x}_i)$ is the unknown function at point $\mathbf{x}_i$. The local risk $\mathcal{R}(\widehat{u}_i, u(\mathbf{x}_i))$ is defined at each point $\mathbf{x}_i$ and then differs from usual performance measures that integrate errors on the whole image. A local adaptation of the neighborhood is more appropriate to improve the estimator in the vicinity of discontinuities. In what follows, we aim at estimating the neighborhood $\Delta(\mathbf{x}_i)$ by minimizing $\mathcal{R}(\widehat{u}_i, u(\mathbf{x}_i))$.

The risk $\mathcal{R}(\widehat{u}_i, u(\mathbf{x}_i))$ can be decomposed into two terms: squared bias $b_i^2$ and variance $v_i^2$ as

$$\mathcal{R}(\widehat{u}(x), u(x)) = \underbrace{(\mathbb{E}[\widehat{u}(x) - u(x)])^2}_{b^2(x)} + \underbrace{\mathbb{E}[(\widehat{u}(x) - \mathbb{E}[\widehat{u}(x)])^2]}_{v^2(x)}.$$

(13)

The variance can be estimated using expression (9), but the bias term $b_i$ depends on the unknown function $u(\mathbf{x}_i)$ and is thus unobservable. However, assuming minimal properties about the unknown function $u$, we can propose an upper bound for the squared bias term and derive the following property for the optimal estimator $\widehat{u}_i^*$:

$$\frac{(b_i^*)^2}{(v_i^*)^2} = \frac{d}{2} \triangleq \gamma^2,$$

(14)

where $d$ is the dimension of the space-time domain. Expression (14) does not depend on image regularity and further is image-dependent [15]. Following the Lepskii's principle [16], we exploit this property to minimize the $L_2$ risk $\mathcal{R}(\widehat{u}_i, u(\mathbf{x}_i))$. This idea is to design a sequence of increasing radius $(h_{i,n})_{n \in [0,N[}$ for the neighborhood $\Delta(\mathbf{x}_i) : \mathcal{H}_i = \{h_{i,n}, n \in [0,N[: h_{i,n-1} \leq h_{i,n}\}$. Assuming that the variance $v_{i,n}^2$ is a decreasing function of $n$ (this is especially true when the number of point in the neighborhood increase quickly), the number of samples taken into account is progressively increased to reduce the estimator variance while controlling the estimator bias. Formally, the so-called "bias-variance trade-off" corresponds to the following inequality:

$$h_i^* = \sup_{h_{i,n} \in \mathcal{H}_i} \{|b_{i,n}| \leq \gamma v_{i,n}\}.$$

(15)

This stepwise procedure will allow us to select among a predefined list of neighborhoods $\{\Delta_{i,n}, n \in [0,N[\}$, the neighborhood that best minimizes the local quadratic risk. Since $b_{i,h}$ is unknown, we consider instead a weaker "oracle" to detect the optimal window for smoothing [17]:

$$h_i^* = \sup_{h_{i,n} \in \mathcal{H}_i} \{n' < n \ : \ |\widehat{u}_{i,n} - \widehat{u}_{i,n'}| \leq \rho v_{i,n'}\}.$$

(16)

where $\rho = 2\sqrt{2}$ is a positive constant. The design of a sequence of increasing neighborhoods is now required for estimation. However, in the case of image sequence, the relationship between the temporal and spatial dimensions is related to the object size and movement, which are both unknown. Accordingly, the space and time extents of neighborhoods should be considered independently. For this reason, we decide to increase alternatively the size of the neighborhood using two distinct radii. We note respectively $h_s$ and $h_t$ the spatial and temporal neighborhoods which can vary from one point to another.

| Sequences | $L_1$ | | $L_2$ | | $t_e$ |
|:---:|:---:|:---:|:---:|:---:|:---:|
| | mean | std | mean | std | |
| N | 4.39 | $6 \cdot 10^{-3}$ | 35.0 | $12 \cdot 10^{-3}$ | |
| A | 1.56 | $16 \cdot 10^{-3}$ | 2.94 | $28 \cdot 10^{-3}$ | 65 min |
| B | 1.96 | $17 \cdot 10^{-3}$ | 3.78 | $25 \cdot 10^{-3}$ | 55 min |
| C | 1.65 | $14 \cdot 10^{-3}$ | 3.01 | $24 \cdot 10^{-3}$ | 28 min |

**Table 1**. Influence of the variance stabilization transform and of the adjacent temporal frames on the error. $L_1$ and $L_2$ norms are used to measure the performance of the denoising procedure. The mean and standard deviation wrt time sequence are reported. The time computings $t_e$ for each experiment is also given for the following sequences: N: noisy sequence; A: 3D+t, Gaussian and Poisson noise; B: 3D+t, Gaussian noise; C: 3D, Poisson and Gaussian noise.

It is worth noting that, unlike [17], the sequence of neighborhoods is not known in advance since we consider two parameters $h_s$ and $h_t$. In our experiments, we use a dyadic scale in space and a linear scale in time.

## 3. EXPERIMENTS

### 3.1. Simulated image sequence

In order to analyze performance of the proposed method, we consider three experiments applied on a simulated sequence of 50 images of size $256 \times 256 \times 10$. We have evaluated both the variance stabilization procedure and the 3D+t denoising procedure through three experiments reported in Table 1. For all these three experiments the same parameters are used and $5 \times 5 \times 5$ patches are used. Results confirms that 3D+t denoising and appropriate noise modelling are require to achieve high performances. We can also notice that 3D+t processing implies an additional computationnal load.

### 3.2. Real video-microscopy data

In order to illustrate the performance of the proposed approche, we have also conducted experiments on real image sequences obtained using a spinning disk system. An example is reported on figure 1. Qualitative comparisons with median filtering commonly used to pre-process noisy video-microscopy image sequence show that the proposed method better adapt to the geometry of images.

Finally, we have conducted a last experiment on fixed samples comparing the improvement brought by the proposed 3D denoising procedure when the exposure time is reduced from 500ms (taken as a reference) to 30ms.

## 4. CONCLUSION

In this paper, we have presented a method for denoising 3D+t video-microscopy image sequences able to take into account the space-time content and noise characteristics. The proposed method relies on the averaging of similar patches lying inside an adpated space-time neighborhood. Optimal adaptation is performed by using the bias-variance tradeoff for the designed estimator. Experiments on both synthetic data and real image sequences shows that our approach outperform standard methods and will allow to reduce illumination and exposure time.
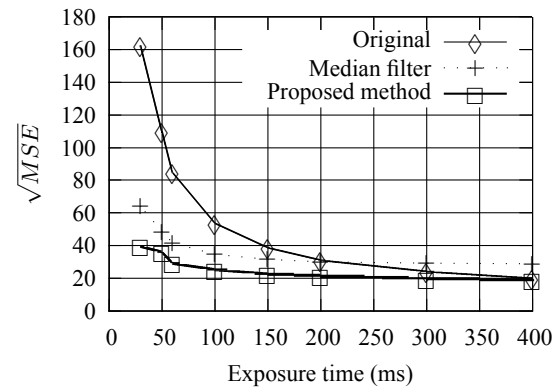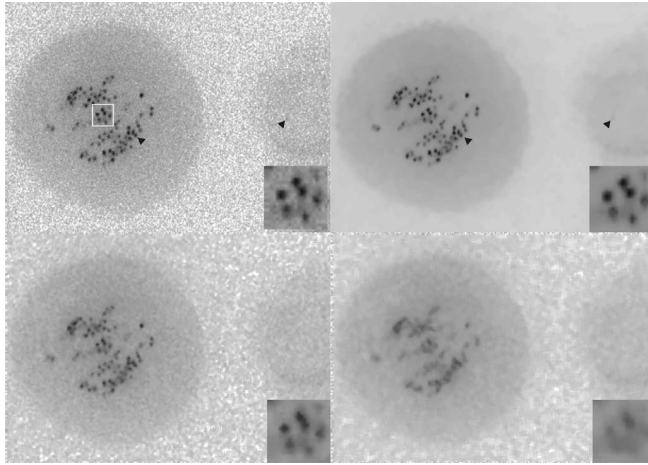
**Fig. 1**. On the left: GFP-Nup 133-Cter + GFP-CENPA, stable HeLa cell line. 3D MIP reconstruction of a stack composed of 8 planes (Dz=2$\mu$m). Stack extracted from a 3D+time serie acquired with a spinning disk microscope, 100x, 1.4NA (M. Zuccolo, V. Doye, UMR144-CNRS, Institut Curie). On the upper-right corner is represented the result of the proposed method while the second row contains the result obtained repectively using $3 \times 3$ and $5 \times 5$ median filtering. Extracted zooms shows that the propose approch outperform tradionnal approches used in fluorescence microscopy. On the right, the square root of the mean squared error is plotted against exposure time showing the improvement of the filtering in the case of a fixed sample.

## 5. REFERENCES

[1] J. C. Brailean, R. P. Kleihorst, S. Efstratiadis, A. K. Katsagge-los, and R. L. Lagendijk, "Noise reduction filters for dynamic image sequences: A review," *Proc. of the IEEE*, vol. 83, no. 9, pp. 1272–1291, Sept. 1995.

[2] A. Kuznetsov, V. P. Bindokas, J. D. Marks, and L. H. Philipson, "FRET-based voltage probes for confocal imaging: membrane potential oscillations throughout pancreatic islets," *American Journal Physiology - Cell Physiology*, vol. 289, no. 1, pp. 224–229, Mar. 2005.

[3] N. Rajpoot, Zhen Yao, and R. Wilson, "Adaptive wavelet restoration of noisy video sequences," in *Proc. of ICIP'2004*, Singapore, 2004, pp. 957–960.

[4] F. Shi and I. W. Selesnick, "Video denoising using oriented complex wavelet transforms," in *Proc. of the IEEE ICASSP'04*, Montreal, May 2004, vol. 2, pp. 949–952.

[5] F. Dekeyser, P. Bouthemy, and P. Pérez, "Spatio-temporal Wiener filtering of image sequences using a parametric motion model," in *Proc. of ICIP'2000*, Vancouver, Canada, Sept. 2000, pp. 208–211.

[6] D. Uttenweiler, C. Weber, B. Jähne, R.H. Fink, and H. Scharr, "Spatio-temporal anisotropic diffusion filtering to improve signal to noise ratios and object restoration in fluoresence microscopic image sequences," *Journal of Biomedical Optics*, vol. 8, no. 1, pp. 40–47, Jan. 2003.

[7] A. Buades, B. Coll, and J.-M. Morel, "Denoising image sequences does not require motion estimation," in *IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2005*, 2005.

[8] K. Dabov, V. Foi, A. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3d transform-domain collaborative filtering," *IEEE Trans. on IP*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[9] J. Mairal, G. Sapiro, and M. Elad, "Learning multiscale sparse representation for image and video restoration," *SIAM*, 2007, submitted.

[10] R. M. Willett and R. D. Nowak, "Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging," *IEEE Trans. on TMI*, vol. 22, pp. 332–350, Mar. 2003.

[11] F. J. Anscombe, "The transformation of Poisson, binomial and negative-binomial data," *Biometrika*, vol. 35, pp. 246–254, 1948.

[12] J.-L. Starck, F. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis, the Multiscale Approach*, Cambridge University Press, 2000.

[13] C. L. Chan, A. K. Katsagellos, and A. V. Sahakian, "Image sequence filtering in quantum-limited noise with application to low-dose fluoroscopy," *IEEE Trans. on TMI*, vol. 12, no. 3, pp. 610–621, Sept. 1993.

[14] T. Gasser, L. Sroka, and C. Jennen-Steinmetz, "Residual variance and residual pattern in nonlinear regression," *Biometrika*, vol. 73, pp. 625–633, 1986.

[15] L. Stankovic, "Performance analysis of the adaptive algorithm for bias-to-variance tradeoff," *IEEE Trans. on SP*, vol. 52, no. 5, pp. 1228–1234, May 2004.

[16] O. Lepski, "Asymptotically minimax adaptive estimation 1: upper bounds," *SIAM Journal Theory of Probability and Application*, vol. 36, no. 4, pp. 654–659, 1991.

[17] C. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch-based image denoising," *IEEE Trans. on IP*, vol. 15, no. 10, pp. 2866–2878, 2006.